



## Article

# Investigating Dual-Source Satellite Image Data and ALS Data for Estimating Aboveground Biomass

Wen Fan <sup>1</sup>, Jiaojiao Tian <sup>1,2</sup>, Thomas Knoke <sup>1</sup>, Bisheng Yang <sup>3,4,\*</sup>, Fuxun Liang <sup>3,4</sup> and Zhen Dong <sup>3,4</sup>

<sup>1</sup> Institute of Forest Management, Department of Life Science Systems, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany; wendy.fan@tum.de (W.F.); jiaojiao.tian@tum.de (J.T.); knoke@tum.de (T.K.)

<sup>2</sup> Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany

<sup>3</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; liangfuxun@whu.edu.cn (F.L.); dongzhenwhu@whu.edu.cn (Z.D.)

<sup>4</sup> Engineering Research Center for Spatio-Temporal Data Smart Acquisition and Application, Ministry of Education of China, Wuhan University, Wuhan 430079, China

\* Correspondence: bshyang@whu.edu.cn

**Abstract:** Accurate estimation of above-ground biomass (AGB) in forested areas is essential for studying forest ecological functions, surface carbon cycling, and global carbon balance. Over the past decade, models that harness the distinct features of multi-source remote sensing observations for estimating AGB have gained significant popularity. It is worth exploring the differences in model performance by using simple and fused data. Additionally, quantitative estimation of the impact of high-cost laser point clouds on satellite imagery of varying costs remains largely unexplored. To address these challenges, model performance and cost must be considered comprehensively. We propose a comprehensive assessment based on three perspectives (i.e., performance, potential and limitations) for four typical AGB-estimation models. First, different variables are extracted from the multi-source and multi-resolution data. Subsequently, the performance of four regression methods is tested for AGB estimation with diverse indicator combinations. Experimental results prove that the combination of multi-source data provides a highly accurate AGB regression model. The proposed regression and variables rating approaches can flexibly integrate other data sources for modeling. Furthermore, the data cost is discussed against the AGB model performance. Our study demonstrates the potential of using low-cost satellite data to provide a rough AGB estimation for larger areas, which can allow different remote sensing data to meet different needs of forest management decisions.

**Keywords:** airborne laser scanning (ALS); satellite imagery; forest above-ground biomass; multi-source data; multiple regression



**Citation:** Fan, W.; Tian, J.; Knoke, T.; Yang, B.; Liang, F.; Dong, Z. Investigating Dual-Source Satellite Image Data and ALS Data for Estimating Aboveground Biomass. *Remote Sens.* **2024**, *16*, 1804. <https://doi.org/10.3390/rs16101804>

Academic Editor: Henning Buddenbaum

Received: 18 March 2024

Revised: 6 May 2024

Accepted: 17 May 2024

Published: 19 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Carbon peaking and carbon neutrality goals are currently popular topics all over the world [1]. Forests account for a high proportion of global above-ground carbon stocks [2]. An accurate AGB estimation is the data basis for studying forest carbon stocks and investigating their impact on climate change and ecosystem functions. However, traditional forest inventory for biomass estimation includes inevitable drawbacks when taking large-scale measurements [3], such as great time consumption and high costs, making it difficult to popularize. Remote sensing data contribute to obtaining spatially contiguous high-precision data [4,5]. Among these, the cost and performance of the AGB estimation model are the main factors determining the selection of remote sensing data sources.

The observational advantages of remote sensing, such as wide coverage and fast operation, compensate for the shortcomings of traditional methods of measuring forest biomass. Current remote sensing methods for estimating AGB mainly utilize optical remote sensing and light detection and ranging (LiDAR). Optical images contain spectral

information of the forest on a large scale. The spectral information obtained from satellite images has a good correlation with AGB, representing its potential to inform predictors in AGB regression. However, if the tree density is high it can be difficult to quantify a reliable statistical relationship between spectral information and biomass [6]. Unlike optical image systems, LiDAR systems capture the vertical distance between the sensor and the object. The 3D structural data of vegetation generated by LiDAR has been adopted in many applications [7]. Synthetic Aperture Radar (SAR) at lower frequencies (L and P bands) has the ability to penetrate deep into the forest and interact with the trees, providing a better correlation with AGB [8]. However, it has stronger ground effects than LiDAR. Owing to their distinct features, optical images and laser scanning point clouds can be combined to monitor forest AGB changes.

The synthesis of multi-source features for modeling AGB has been widely used for AGB estimation [9]. For example, optical image data and ALS data [10]. de Almeida et al. [4] explore the combination of ALS and hyperspectral data, which showed significantly improved performance in Amazon AGB estimation. It has been proved that remote sensing data sources (ALS, Airborne hyperspectral imagery, or their combination) have a greater impact on modeling outcomes than regression methods. In addition, there are many challenges in the selection of suitable metrics, as well as many different models [9]. In contrast to deep learning methods, various machine learning approaches have been explored to investigate the importance of different variables in the estimation of AGB. The choice of parametric and non-parametric statistical models is crucial for AGB estimation, as the performance of models varies significantly in different modeling methods [3]. Previous studies have claimed that their developed AGB estimation models can achieve satisfactory results using variable selection, different regression models, and multi-source and multi-resolution data. However, to the best of our knowledge, there has been little prior research on the fusion of image data with different resolutions and ALS data in AGB modeling.

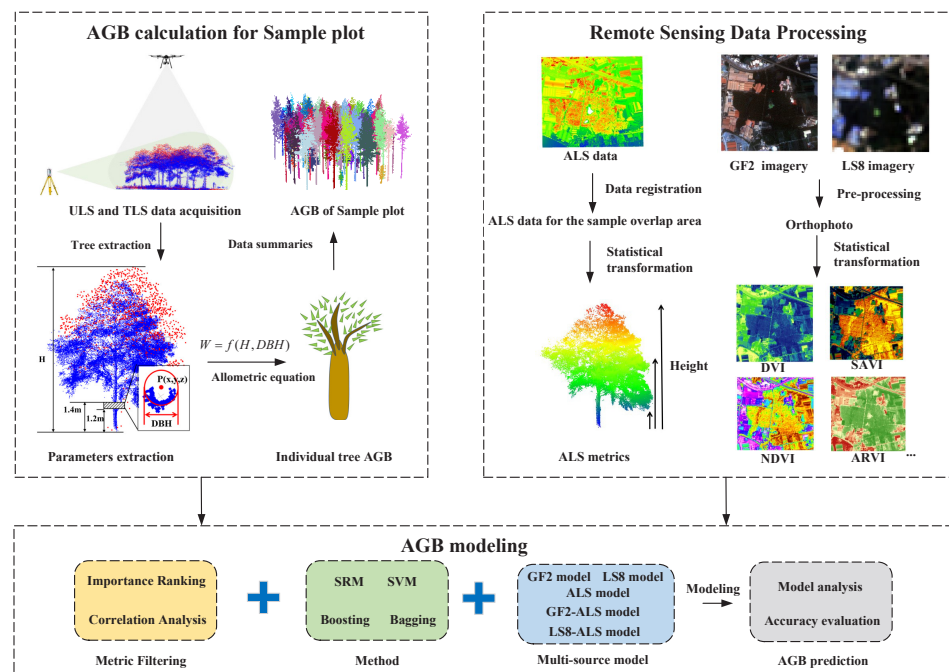
To address these knowledge gaps, our work combines medium- and high-resolution satellite data with airborne point clouds. The impact of different resolutions of satellite imagery on modeling performance is investigated. Satellites data have large coverage and are publicly available, while ALS data can capture 3D tree morphology information but usually cover a smaller area owing to budgetary limitations. This study proposes a framework to leverage the different advantages of satellite imagery and ALS data in estimating aboveground biomass. Specifically, instead of developing a new algorithm, we assess several regression models using dual-source data. This assessment helps us to identify the best regression model based on model performance and cost, which is then used to estimate the aboveground biomass over large areas. AGB modeling is explored in this paper to answer the following questions: (1) What impact does the integration of ALS data variables and satellite imagery variables have on the precision and bias of AGB estimations? (2) To what extent do the different regression methods affect the accuracy and bias of AGB estimation? and (3) How does the spatial resolution of satellite imagery influence the estimation of AGB?

The remainder of this paper is organized as follows. Section 2 describes the multi-source data, the extracted variables, and the methods used to explore the AGB regression modeling for forested areas. First, the vertical structure of the forest sample plots is obtained by applying terrestrial laser scanning (TLS) and unmanned aerial vehicle laser scanning (ULS) data on field plots. The extracted diameter at breast height (DBH) and tree height of individual trees are then used to estimate plot-level AGB through allometric model equations to obtain the AGB as a dependent variable. Independent variables are extracted from the ALS data and images with different resolutions in order to explore the correlation between the different variables and forest AGB. Section 3 describes the experimental results and evaluation. A discussion is presented in Section 4, in which AGB models built by the different methods and data sources are compared and analyzed. Finally, the paper is summarized in Section 5. This study explores the performance of AGB modeling with a

fusion data source and investigates the impact of integrating spectral information with different cost considerations into high-cost ALS data.

## 2. Materials and Methods

As illustrated in Figure 1, the proposed method includes the following main steps: (1) tree parameter and AGB calculation at the sample plot level; (2) extraction and selection of variables from satellite images and ALS data; and (3) AGB regression model construction and evaluation using four regression methods and multi-source data.



**Figure 1.** The workflow of the proposed method.

### 2.1. Study Areas and Datasets

Guangxi province, China, has a subtropical monsoon climate with abundant rainfall and heat. Most of the forest plots are planted trees. In this study, four forested areas in Guangxi province were selected, as listed in Table 1 and Figure 2. In the experiment, a total of 68 plots of  $15 \times 15$  m were set up. The experimental data included ULS data, TLS and ALS data, Landsat 8 (LS8) data, and Gaofen2 (GF2) image data. The ULS and TLS data were collected in June 2020. The ALS data and satellite images were collected in 2019 and 2020, respectively.

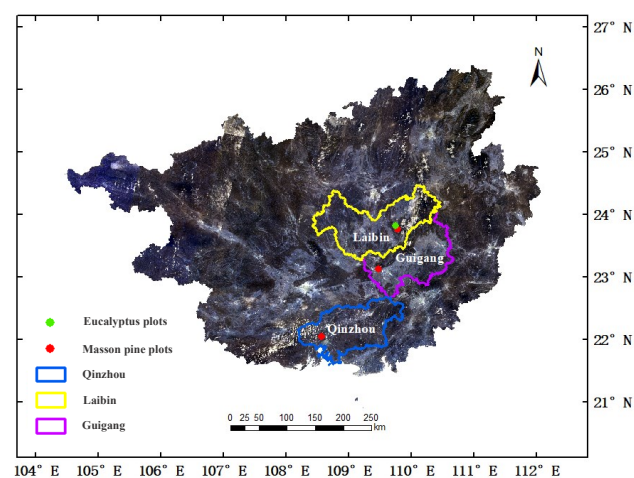
The ULS, airborne, and terrestrial LiDAR data allowed for complete scanning of the internal and upper structure of the forest. An Austria *RieglVZ-400* 3D laser scanning system, which scans 122,000 points per second, was used for the TLS data acquisition in the experiment. A Netherlands *Aerialtronics Altura AT8* ULS, which has a flight height of about 50–70 m above the ground, was used for the unmanned aerial vehicle (UAV) data acquisition. The ALS data were acquired using a fixed-wing manned *P750* aircraft equipped with an Austria *RIGEL-VQ-1560i* airborne LiDAR scanning system. The ALS system was used at an average flight altitude of 2500 m above ground, with a full area point density of over three points/m<sup>2</sup>. The ALS data were preprocessed to produce a digital terrain model (DTM) and a normalized point cloud.

For the sample plots, the GF2 satellite images were downloaded from the China Center for Resources Satellite Data and Application, which provides cloud-free images at the sub-meter level. For the sample plots, the GF2 satellite provides cloud-free images at the sub-meter level. The images have four spectral bands with a spatial resolution of 4 m (wavelength 0.45–0.9  $\mu\text{m}$ ) and one panchromatic band with a spatial resolution of 1 m.

Similarly, the LS8 images have nine multispectral bands with a spatial resolution of 30 m along with a panchromatic band with a resolution of 15 m. Pansharpening was performed on the GF2 and LS8 data using the software Esri *ENVI*5.3 from Davis, CA, USA. After pansharpening, the images were cut and mosaicked, then the preprocessed satellite images of the study area were obtained.

**Table 1.** Study area description.

Area	Tree Species	Plots
Guigang	Masson pine	20
Laibin-1	Masson pine	12
Laibin-2	Eucalyptus	16
Laibin-2	Masson pine	20



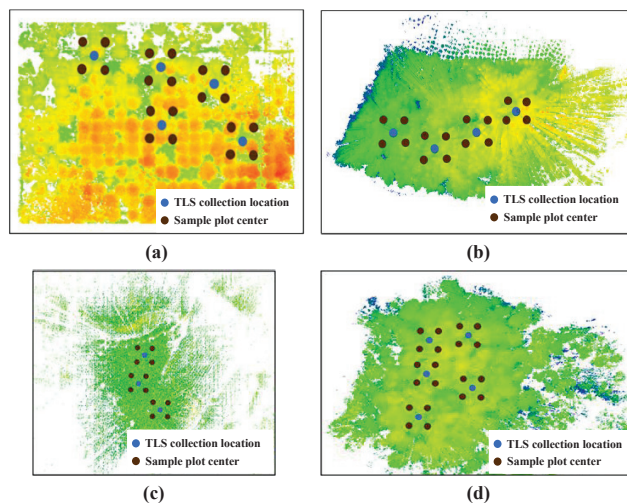
**Figure 2.** The study areas in Guangxi Province, China.

## 2.2. Construction of Field Plots

According to Table 1, all TLS and ULS data in four forests in three cities were collected. The distribution of sample plots is shown in Figure 3. The collected point clouds can be used to segment and calculate individual trees. The accuracy of the point cloud method for obtaining tree parameters has been verified in [11]. The AGB of an individual tree is the sum of the parts, including the stems, branches, and leaves. The AGB sample plots were assessed based on the existing allometric equations [12], as shown in Table 2. TLS obtains rich information on the bottom part of the tree, based on which the *DBH* can be accurately calculated. ULS is able to capture information on the top of the tree, the ground height, and the highest point of the tree, which are important for correctly calculating *H*. The combination of TLS and ULS data can be used to obtain complete information on the sample plots. First, the collected multi-station TLS point clouds were co-registered to the same coordinate systems, then the TLS and ULS data were co-registered using the Random Sample Consensus (RANSAC) method [13]. The average registration residual for the TLS-to-TLS scenario was 0.049 m, while for the ULS-to-ALS scenario it was 0.299 m. The fused point cloud was then used to automatically perform an accurate individual tree extraction [14]. In addition, we used the TLS-to-ALS registration from [15]. The average registration residual for the TLS-to-ALS scenario was 0.049 m. The conversion of ALS data to 1 m resolution imagery allowed for geographic registration of the ALS and GF2 imagery using the geographic registration method in *ENVI* 5.3. Satellite image plots with the same geographic extent were obtained from TLS point cloud plots.

**Table 2.** Allometric equations used for AGB estimation from [12].

Tree Species	Part	$W = a(DBH)^b$		$r^2$
		$a$	$b$	
Masson pine	Stems	0.1615	2.2989	0.88
	Branches	0.0763	1.8402	0.88
	Leaves	0.3248	0.935	0.65
Eucalyptus	Stems	0.0702	2.5253	0.84
	Branches	0.0452	1.794	0.74
	Leaves	0.7392	0.2565	0.02

**Figure 3.** Distribution of sample sites and TLS collection locations: (a) Masson pine forest in Guigang municipality; (b) Masson pine forest in Laibin municipality; (c) Eucalyptus forest in Laibin municipality; and (d) Masson pine forest in Qinzhou municipality.

### 2.3. Variable Extraction and Selection

#### 2.3.1. Satellite Image Processing and Metric Extraction

Several imagery spectrum variables or transformation results have been proposed as potential AGB predictors. Based on the LS8 and GF2 images, some of the optical image variables were obtained by principal component analysis (PCA) transformation, minimum noise fraction (MNF) transformation, and transformation by various vegetation indices [5]. Descriptions of the variables and the calculation equations are provided in Table 3. All the satellite image variables were obtained at the plot level.

For GF2 data and LS8 data, the raw bands have different spectral characteristics in terms of the absorption and reflectance of vegetation, and are used as candidate variables, as shown in Table 3. Here,  $\rho_B$ ,  $\rho_G$ ,  $\rho_R$ , and  $\rho_{NIR}$  correspond to Bands 1 to 4, respectively. PCA is able to extract the most important components of the image, and is commonly used to remove noise from satellite data. The first three principal components are not correlated with each other after transformation. The data variance of the first principal component (PC1) is the largest, then those of PC2 and PC3 decrease in turn. The retained three MNF components arrange the components from largest to smallest according to the signal-to-noise ratio. In addition, the raw bands can be combined to calculate vegetation indices in order to highlight the vegetation. For example, the NDVI is used to provide an indication of the health and growth of the vegetation; the EVI represents an improvement over NDVI in terms of a decoupling the canopy background signal and reducing atmospheric influences. The RVI and DVI reflect the difference between the reflection of vegetation in the visible and near-infrared bands. The SAVI and MSAVI are used to reduce the effect of the soil. The ARVI is insensitive to aerosols, and is particularly suitable for monitoring areas with high atmospheric aerosol levels [16].



**Table 3.** Metrics calculated from the satellite image data.

Abbr.	Description	Equation	Reference
$\rho_R, \rho_G, \rho_B, \rho_{NIR}$	Raw optical image bands		
PC1, PC2, PC3	Extraction of the three band of PCA		
MNF1, MNF2, MNF3	Extraction of the three band of MNF transformation		
EVI	Enhanced vegetation index	$EVI = 2.5 * (NIR - R) / (NIR + 6.0 * R - 7.5 * B + 1.0)$	[17]
NDVI	Normalized difference vegetation index	$NDVI = (NIR - R) / (NIR + R)$	[18]
RVI	Ratio vegetation index	$RVI = NIR / R$	[19]
DVI	Differential vegetation index	$DVI = NIR - R$	[20]
SAVI	Soil-adjusted vegetation index	$SAVI = (1.0 + 0.5) * (NIR - R) / (NIR + R + 0.5)$	[21]
MSAVI	Modified soil-adjusted vegetation index	$MSAVI = NIR + 0.5 * \sqrt{(NIR + 0.5)^2 - 2 * (NIR - R)}$	[22]
ARVI	Atmospherically resistant vegetation index	$ARVI = (NIR - (2.0 * R - B)) / (NIR + (2.0 * R - B))$	[23]

### 2.3.2. ALS Data Processing and Variable Extraction

A range of ALS variables are used to characterize the forest canopy and vertical structures, including height percentile variables, canopy cover variables, and tree height variables at the plot level. In this case, to construct the height percentile variables within a sample plot, the height at a given percentile is extracted from the normalized point cloud. The height percentile variables show the vertical division of the forest. All the extracted ALS variables are listed in Table 4.

**Table 4.** Summary of the ALS metrics.

Abbr.	Description
C	Canopy cover.
$h_{max}$	Maximum height of first returns above 2 m
$h_{mean}$	Mean height of first returns above 2 m
$h_{std}$	Height standard deviation of first returns above 2 m.
$h_{cv}$	Height coefficient of variation (%) of first returns above 2 m.
$h_{25}, h_{50}, h_{60}, h_{75}, h_{95}$	$X_{th}$ (25, 50, 60, 70, 75 or 95th) percentile of height distribution of first returns above 2 m
FR	Percentage of first return points to total points
$t_{max}$	Maximum of ALS tree height
$t_{mean}$	Average of ALS tree height
$t_{min}$	Minimum of ALS tree height

In general, the first return points of ALS point clouds are the canopy points in the forest, while the last return points are the ground points. The percentage of first return points to total points (FR) variable describes the proportion of canopy points to the total number of points in the forest. Further extraction of the canopy height variables can describe the vertical distribution and variability of the forest canopy. Canopy cover is defined as the percentage of vegetation returns to the total number of returns, which describes the planting density of trees at the horizontal level.

### 2.3.3. Selection for Satellite Imagery and ALS Data Variables

As the extracted variables contain redundant information, variable selection is a key issue. Importance ranking of variables and Pearson correlation analysis was used to filter the variables. Finally, suitable optical image and ALS data variables, along with their combinations, were added to the regression model.

After extracting variables from the preprocessed GF2, LS8, and ALS data, the random forest (RF) importance ranking of the variables was calculated. The change in the mean squared error (%IncMSE) can be considered as the contribution of the variable to the AGB prediction and used to assess the importance of specific variables. In general, the %IncMSE is used to interpret the decrease in the precision of the AGB prediction when the variable is removed. When %IncMSE is higher than 4%, the variable is retained. Meanwhile, the correlation between the extracted variables was calculated using Pearson correlation analysis. In this paper, 0.7 was considered as the assumed threshold for a high correlation coefficient. If the correlation coefficient exceeds 0.7, only the more important variables are

kept. In this way, those variables with high importance and low correlation are selected for modeling.

#### 2.4. Regression Models

The regression methods considered in this study included both parametric and non-parametric methods, namely, the stepwise regression method (SRM), support vector machine (SVM), boosting tree, and bagging tree approaches. Ten-fold cross-validation [24] was employed to prevent overfitting of the models using all 68 plots.

Multiple variable regression analysis is a commonly used method in biomass modeling [25]. In this study, the SRM method was used to describe the relationship between the independent variables and the dependent variable (AGB). When a new dependent variable was introduced, the original variables were tested one at a time. Variables were retained ( $p < 0.05$ ) or removed depending on their significance level. The independent variables from Tables 3 and 4 were filtered via SRM to build the final model, as shown in Equation (1):

$$AGB = a_1 * metric_1 + a_2 * metric_2 + \dots + a_n * metric_n + b \quad (1)$$

where  $b$  is an intercept and  $a_n$  is the parameter for  $variable_n$  fitting in the SRM. For metrics,  $variable_n$  is the one in Table 3 or Table 4 from the images and ALS data retained after feature selection and the SRM.

The other three methods (SVM, boosting tree, and bagging tree) are all non-parametric methods which use default parameters. The same type of learner was used in both the bagging tree and boosting tree methods; the differences are shown in terms of sample selection, sample weight, prediction function, and parallel computation. SVM is a dichotomous classification method that separates two categories by seeking an optimal decision boundary at a maximum interval.

The precision and bias vary depending on the different regression models. To assess the results of the different regression models, the coefficient of determination ( $R^2$ ), root-mean-square error (RMSE), and bias were calculated in order to evaluate the performance of the models. The definitions for the calculation of each indicator are as follows:

$$R^2 = 1 - (\sum_{i=1}^n (y_i - \tilde{y}_i)^2) / (\sum_{i=1}^n (y_i - \bar{y}_l)^2) \quad (2)$$

$$RMSE = \sqrt{(\sum_{i=1}^n (y_i - \tilde{y}_i)^2) / n} \quad (3)$$

$$bias = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 / (n - 1) \quad (4)$$

where  $n$  is the number of samples,  $y_i$  is the ground truth of the  $i^{th}$  sample plot,  $\tilde{y}_i$  is the predicted value for the  $i^{th}$  sample plot, and  $\bar{y}_l$  indicates the mean of the ground truth of the sample plots.

### 3. Experiments Results

#### 3.1. Variables Selection Results

Essentially, a simple model should be more stable and transferable than a complex one. Therefore, variable selection is necessary before employing regression models for AGB, as the complexity of the regression models depends on the number of input variables. Potential variables were further selected from Tables 3 and 4 for AGB regression based on the RF importance ranking and Pearson correlation analysis.

Concerning the outcomes of variable selection, the chosen variables for GF2 data include  $PC3$ ,  $EVI$ ,  $MSAVI$ ,  $PC1$ , and  $RVI$ . In the case of LS8 data, the selected variables encompass  $DVI$ ,  $EVI$ ,  $MNF2$ ,  $MNF3$ , and  $PC3$ . For ALS data, the selected variables encompass  $h_{cv}$ ,  $t_{max}$ ,  $FR$ ,  $h_{mean}$ ,  $t_{min}$ ,  $C$ , and  $h_{max}$ . Additionally, the GF2-ALS model incorporates  $h_{cv}$ ,  $PC3$ ,  $t_{max}$ ,  $h_{mean}$ ,  $EVI$ ,  $RVI$ , and  $MNF3$ , while the LS8-ALS model integrates  $h_{cv}$ ,  $t_{max}$ ,  $PC2$ ,  $h_{mean}$ ,  $FR$ ,  $EVI$ ,  $DVI$ ,  $t_{min}$ , and  $C$ . The significance rankings and correlations of the designated satellite image indicators, ALS data, and amalgamated data are illustrated in Figures A1–A3, respectively, presented in Appendix A.

### 3.2. Performance of the Four Regression Methods

#### 3.2.1. AGB Prediction Using Only Image Variables

Following variable selection based on their correlations and significance, a performance evaluation of multisource data modeling was conducted. In this paper, the AGB regressions using only the GF2 variables or the LS8 variables are called the GF2 model and LS8 model, respectively. The results of each regression model and their associated statistical performance indicators are shown in Table 5. In the GF2 model depicted in Table 5, the obtained  $R^2$  ranges between 0.54 (boosting) and 0.61 (SVM). For the bias and RMSE, the model performance with the largest difference between the predicted value and the ground truth is the boosting model (bias =  $-4.38$  Mg/ha and RMSE = 44.21 Mg/ha), while the lowest is the SRM (bias =  $-0.0003$  Mg/ha) and SVM (RMSE = 40.72 Mg/ha). The bold text in Table 5 is the best model result in this data source.

**Table 5.** AGB modeling performance of the ten-fold cross-validated regression methods.

$(n = 68)$		$R^2$	RMSE (Mg/ha)	rRMSE	bias (Mg/ha)
GF2 model	SRM	0.57	43.61	29.18%	$-0.0003$
	Boosting	0.54	44.21	29.58%	$-4.38$
	<b>SVM</b>	<b>0.61</b>	<b>40.72</b>	<b>27.24%</b>	$-1.50$
	Bagging	0.59	41.82	27.98%	0.16
LS8 model	SRM	0.44	50.01	33.46%	$-0.0002$
	Boosting	0.49	46.97	31.42%	$-5.81$
	SVM	0.42	49.88	33.38%	$-1.19$
	<b>Bagging</b>	<b>0.52</b>	<b>45.40</b>	<b>30.38%</b>	<b>0.71</b>
ALS model	<b>SRM</b>	<b>0.78</b>	<b>31.85</b>	<b>21.31%</b>	<b>0.00008</b>
	Boosting	0.71	35.23	23.57%	$-5.09$
	SVM	0.67	37.63	25.18%	$-2.00$
	Bagging	0.70	36.04	24.11%	1.01
GF2-ALS model	<b>SRM</b>	<b>0.82</b>	<b>28.85</b>	<b>19.30%</b>	<b>0.0002</b>
	Boosting	0.68	37.18	24.88%	$-4.01$
	SVM	0.66	38.29	25.62%	$-4.05$
	Bagging	0.74	33.69	22.54%	$-1.42$
LS8-ALS model	<b>SRM</b>	<b>0.78</b>	<b>31.85</b>	<b>21.31%</b>	<b>0.00008</b>
	Boosting	0.74	33.69	22.54%	$-6.27$
	SVM	0.65	39.21	26.23%	$-2.48$
	Bagging	0.70	35.93	24.04%	1.22

Figure 4 showcases the cross-validation outcomes pertaining to the AGB predictions derived from the four models compared to the AGB estimates at the sample plot level. Additionally, black linear trendlines and  $R^2$  have been incorporated into Figure 4 to depict the trends for both predicted values and measured data points. As the slope approaches unity and the intercept of the linear trendline diminishes, the determination coefficient increases, signifying an improved fit. The observations in Figure 4 associated with the SVM methods manifest a relatively concentrated distribution in closest proximity to the trendline. Concurrently, Figure 5 elucidates the residual distribution exhibited by the four models from diverse data sources. Verification of residual randomness can be conducted by means of residual-versus-fitted-value plots. The presence and magnitude of heteroskedasticity can be ascertained from these plots. An even distribution of the residuals on both sides of the  $x$ -axis in such plots signifies a limited degree of heteroskedasticity, while a conspicuously non-uniform distribution indicates a pronounced presence of heteroskedasticity. As can be inferred from Figure 5, the SVM method exhibits a reduced number in comparison to the other three methods, along with a more tightly concentrated residual distribution.



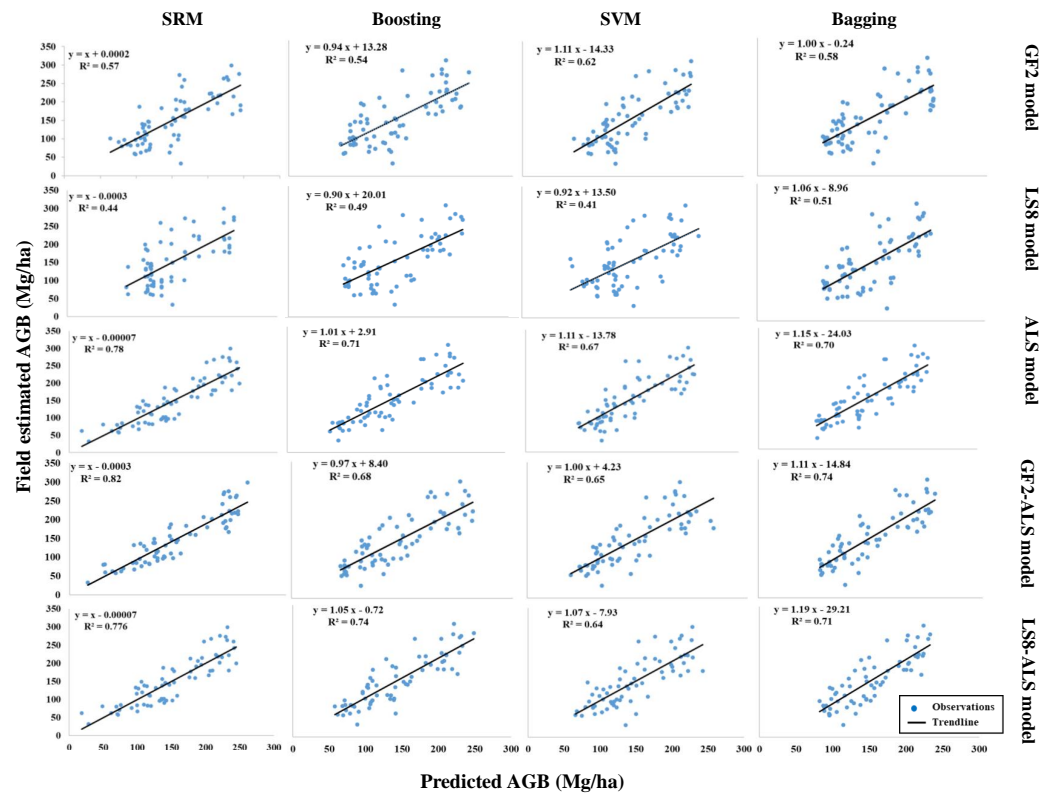


Figure 4. Scatterplots of the predicted and estimated AGB for the four fitting methods.

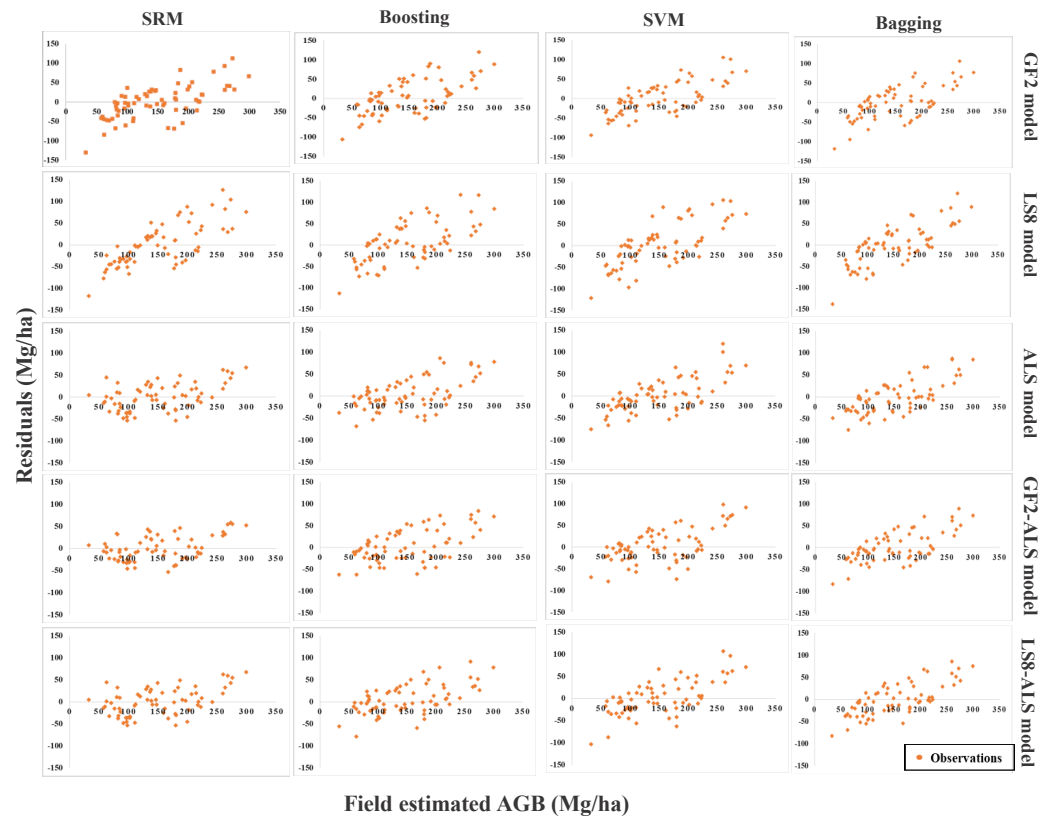


Figure 5. Scatterplots of the residuals and predicted AGB for the four fitting methods.

Similarly, the five extracted LS8 variables were applied to the modeling using the regression method. The results are visually depicted in Figures 4 and 5. Compared with the GF2 model in Table 5, the rRMSE is higher, exceeding 30%. The proportion of explained variation decreases, and the best fitting method is bagging, with an  $R^2$  of 0.52. The lowest  $R^2$  of all the models is 0.42, resulting from the SVM method. The lowest values of RMSE are obtained from the bagging (45.40 Mg/ha) and SRM (50.01 Mg/ha) methods. From Figure 4, the bagging method of the LS8 model exhibits a notably clustered distribution of observations in close proximity to the trendline compared to other methods. Meanwhile, the residual distributions for the various methods demonstrate similarity in Figure 5, where there is an uneven distribution of data on both sides of the  $x$ -axis.

### 3.2.2. AGB Prediction Using Only ALS Variables

The ALS model is a biomass regression model constructed using only ALS data. The results of all four regression models are listed in Table 5 and in Figures 4 and 5. As can be seen from Table 5, among the ALS models regressed by the four methods, SRM achieves the best model performance, with the highest  $R^2$  at 0.78 and the lowest RSME of  $\pm 31.85$  Mg/ha. The boosting method achieves the highest absolute value ( $-5.09$  Mg/ha) for bias, while SRM achieves a low of 0.00008 Mg/ha. From Figures 4 and 5, it can be observed that the SRM approach with the ALS model exhibits exceptional fitting performance and optimal residual distribution performance. From Figure 4, it is evident that the observation points of the SRM method are the closest to the trend line. The fitted line has a slope of 1, and the intercept is minimal. In addition, it is noteworthy that this method tends to underestimate values when AGB estimates are below 50 Mg/ha, as shown in Figure 4. From Figure 5, it is evident that only the SVM method exhibits a noticeable imbalance in the distribution of observation points on both sides of the  $x$ -axis. Meanwhile, the SRM method displays the smallest range of variation along the  $y$ -axis.

### 3.2.3. AGB Prediction Using Multi-Source Data

The GF2-ALS and LS8-ALS models represent the utilization of ALS data in conjunction with GF2 and LS8 data, respectively. For the models built by these two fusion data sets as shown in Table 5, the rRMSE does not exceed 25%, except for the LS8-ALS model using the SVM method.  $R^2$  ranges are 0.66 (SVM) to 0.82 (SRM), 0.65 (SVM) to 0.75 (SRM), and the absolute value of bias ranges from 0.0002 (SRM) to  $-4.05$  Mg/ha (SVM), 0.00008 (SRM) to  $-6.27$  Mg/ha (Boosting), the RMSE spans between  $\pm 28.85$  (SRM) and  $\pm 38.29$  Mg/ha (SVM), between  $\pm 31.85$  Mg/ha and  $\pm 39.20$  Mg/ha (SVM), respectively. For the fusion data, the SRM achieves the best performance on both fused data. For both mixed models, the observations of the SRM method are tightly clustered around the trend line, with the best-fitting trend line. As delineated in Figure 4, when the field-estimated AGB surpasses 250 Mg/ha, it is noteworthy that all models sourced from diverse data origins tend to exhibit an AGB underestimation. Both of these fused data models also show a tendency for underestimation by the SRM method for AGB measurements below 50 Mg/ha. Furthermore, Figure 5 demonstrates the use of the SRM method in the GF2-ALS model results in a more favorable residual distribution compared to the LS8-ALS model. Compared to the other three methods, the residual approach of SRM demonstrates a more even distribution on both sides of the  $x$ -axis, with a narrower range of residual values.

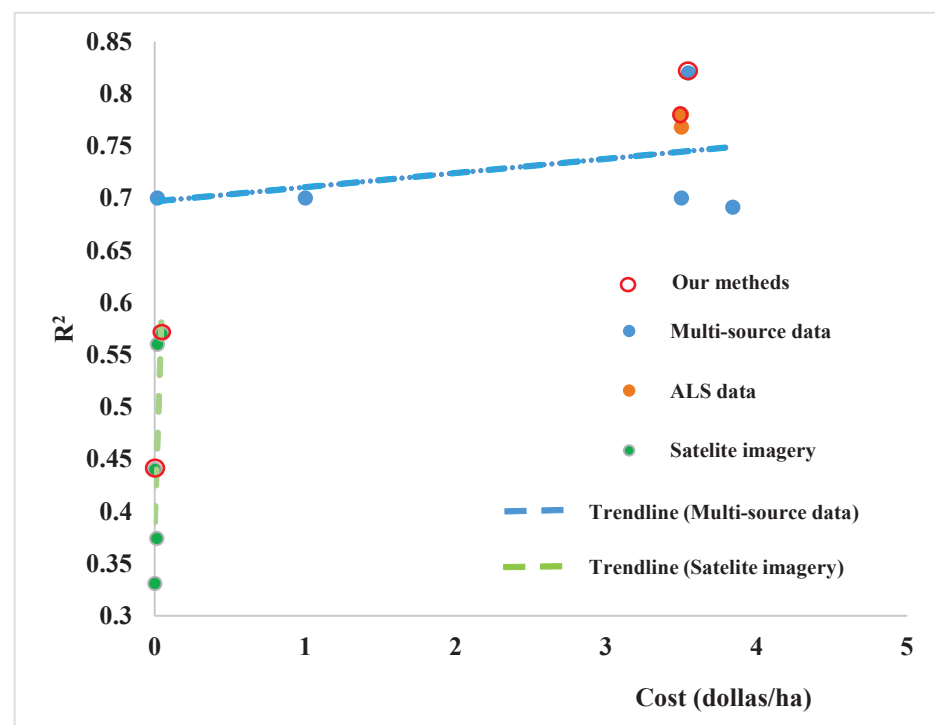
### 3.2.4. Cost and Performance of AGB Modeling

The model performance and costs of the proposed method in this study were compared with existing numbers from the literature [4,9,26,27], with the results shown in Figure 6 and Table 6. As can be seen in Table 6, the source of the cost for satellite imagery per hectare was derived from [28]. The cost per hectare for ALS point clouds represents the current cost incurred for data acquisition. In Figure 6, the data sources in Table 6 are classified into satellite image data, ALS data, and multisource data. The five models mentioned in this

paper are depicted using points in their respective colors, combined with red circles for emphasis.

**Table 6.** Comparison of image data source modeling performance and cost.

Source	Sensors	$R^2$	RMSE (Mg/ha)	rRMSE	Costs (Dollars/ha)	Tree Species
Rana et al. [27]	ALS	0.77	56	0.31	3.50	Mixed forests such as shorea robusta
	RapidEye	0.37	93	0.52	0.01	
	Landsat 5 TM	0.33	96.10	0.54	0	
Han et al. [26]	Gaofen1(GF1)	0.56	19.66	Unknown	0.02	Mix forest such as picea
	GF1 + Sentinel1	0.70	16.26	Unknown	0.02	
de Almeida et al. [4]	Airborne hyperspectral imagery + ALS	0.70	57.70	0.31	3.50	Mixed forest such as palms
Zhang and Shao [9]	WorldView3 + ALS	0.69	26.98	0.44	3.84	Unknown
Yang et al. [29]	UAV + Sentinel2	0.70	70.03	Unknown	1	Mixed forest such as Coniferous
	LS8	0.44	45.40	0.30	0	
This paper	GF2	0.57	40.72	0.27	0.05	Eucalyptus and masson pine
	ALS	0.78	31.85	0.21	3.50	
	GF2 + ALS	0.82	28.85	0.19	3.55	
	LS8 + ALS	0.78	31.85	0.21	3.50	



**Figure 6.** Scatterplot of  $R^2$  and cost. Cost results are for the multispectral products available from the Apollo Mapping archives (Standard Tasking) [28].

## 4. Discussion

### 4.1. Data Source Selection for AGB Modeling

Appropriate choice of data sources for AGB modeling is one of the goals of this paper. In this paper, the performance of ALS data is found to be better for AGB modeling than that of GF2 and LS8 data, as indicated in Table 5. Although good model performance has been achieved in this study, problems with ALS data, such as cost and flight conditions, could affect the further application of ALS in forestry. Satellite images have the advantages of large coverage and frequent monitoring. The issues worth investigating are whether satellite imagery can be used as a single data source for the estimation of AGB, and the impact of resolution improvement on AGB modeling of single data and fused data. We hypothesize that if the image resolution can be further increased, this will provide richer spectral details for the model and may significantly improve the results. In this paper,

GF2 data with 1 m resolution and LS8 data with 30 m spatial resolution were used for regression modeling. The spatial resolution of satellite imagery plays an important role in AGB prediction; all of our models that used GF2 as the input achieved better performance than those using LS8, as shown in Table 5. The experiment results demonstrate that GF2 imagery with 1 m resolution can provide more abundant spatial details than LS8 imagery with 30 m resolution.

LiDAR data are expensive, while optical imagery are susceptible to saturation phenomena. Therefore, the use of fused data is explored in this paper, with the results shown in Table 6 and Figure 6. Rana et al. [27] used ALS data, RapidEye imagery and Landsat 5 TM imagery to predict AGB. RapidEye information at 5 m spatial resolution and Landsat 5 (TM) images at 30 m spatial resolution showed little improvement of the regression results, as shown in Table 6, while the ALS data clearly enhanced AGB regression. In particular, the  $R^2$  value witnessed an improvement of nearly 40%. Han et al. [26] tested combining the LiDAR data with image data, but found no significant effect over the use of the GF1 image data alone when including Sentinel 1 data. This may be due to the roughness of the forest canopy provided by SAR data, which has little effect on AGB prediction. On this basis, our study confirms the reliability of combining satellite imagery and LiDAR point clouds in AGB estimation, which is consistent with previous study results [30]. Compared with the multi-source data modeling results of de Almeida et al. [4] and Han et al. [26], we achieved better AGB estimation results, as illustrated in Figure 6. Therefore, to further improve the precision of ALS estimation of AGB, imagery with a high spatial resolution is needed in order to provide more detailed spatial details.

#### 4.2. Variable Selection for AGB Modeling

Another contribution of this work is the variable selection step using importance ranking and correlation analysis. Remote sensing variables (spectral information, vegetation indices, structural information) are used to correlate with AGB biophysical variables. Variable selection reduces high correlations between two or more predictors and identifies valuable remote sensing variables.

In previous studies, spectral bands, vegetation index, and texture [5] have been the main variables derived for use in AGB estimation in optical remote sensing. Height variables [29] and laser penetration rate [3] are important variables for LiDAR. Based on our exploration of the importance of different variables for AGB modeling,  $PC3$  and  $h_{cv}$  are important and reliable variables with good generalizability. This is consistent with previous research [31]. For satellite images,  $PC2$  and  $PC3$  generated by PCA perform best, while  $PC1$  performs poorly, as shown in the Appendix (see Figure A1). For ALS data,  $h_{cv}$  indicates the degree of dispersion of the data, which is more informative than the numerical variation of the height. The important variables identified in this paper for the estimation of AGB are comparable to the variables used in other research, such as height percentiles [30], height variables [10], and canopy cover [3]. The type and complexity of the forest and the growth of the trees may lead to differences in the structure and spectral information in data from different areas. In future research, more types of variables could be used for estimating AGB; the addition of new variables and groups of multiple variables are topics worth exploring.

#### 4.3. Improvement of AGB Modeling Performance

Model selection has been explored to find the most appropriate regression methods for AGB modeling. de Almeida et al. [4] found that for high spectral resolution data, the choice of regression model has almost no effect. They used six methods, including linear models with (LMR) and without (LM) regularization, Support Vector Regression (SVR), Stochastic Gradient Boosting (SGB), RF, and Cubist (CB) to regress the Brazilian Amazon AGB. Their results were not consistent with previous research findings. They explain that the reasons for the differences could include the number and type of indicators selected as potential input data, the type of vegetation studied, the quality of the field,

and the remote sensing data used to obtain the model. In addition, deep learning (DL) methods have become very popular in recent years [32], although there are still several limitations, such as high data requirements, model complexity, and low interpretability. In this paper, there was no obvious difference between different regression models with the same data source input, except that the SRM model achieved the best performance with the highest  $R^2$  along with the lowest RMSE and rRMSE among the twelve regression models with input. However, SRM is more effective only when the ALS variable is added. In the image data source model, SRM has average performance, as shown in Table 4 and Figure 4. For the model with LiDAR variables, the regression performance of the AGB model constructed by SRM is greatly improved, which has not been found before. SRM adds the ALS variables that are important for AGB modeling and removes the insignificant variables. The ALS variables are more correlated with AGB; therefore, further screening of the variables could significantly improve the performance of the model. Future studies should additionally consider modeling errors and the number of increasing plots.

#### 4.4. Cost vs. Performance

Cost is a critical factor in practical applications, as solutions that balance accuracy and minimum cost are often preferred. While airborne LiDAR provides high accuracy, its ability to improve the accuracy of images at different resolutions in the same scene remains unknown. Therefore, the key issue is that estimating different costs affects the accuracy of biomass estimation. All data sources have been categorized into satellite imagery, airborne LiDAR, and fused data in Figure 6. A clear and intuitive representation of their respective performance and cost is shown in Figure 6 and Table 6. In general, for satellite imagery, higher data quality typically corresponds to an increase in the cost of the image. From Figure 6, it can be observed that as data costs gradually increase, the performance of models improves as well. The optimal results from free remote sensing imagery are obtained through the LS8 model employed in this study, whereas the optimal model performance for a fee is achieved using the GF2 model. Nevertheless, transitioning from LS8 data to GF1 data leads to a cost increase of USD 0.02 per hectare, accompanied by an improvement in  $R^2$  of 0.12. Shifting from the use of GF1 imagery to GF2 imagery results in a cost increase of USD 0.03 per hectare, with an  $R^2$  improvement of 0.01. For fused data, although our models accomplished the best model performance, there is an issue to consider. The cost of fused data using SAR data and GF1 imagery is reduced by USD 3.53 per hectare compared to the GF2-ALS model, with an  $R^2$  decrease of 0.12, though it still achieves an  $R^2$  of 0.7. Furthermore, the free Sentinel-1 imagery is more cost effective compared to ALS data. The process of choosing the most appropriate data source guided by a combination of model performance and cost considerations can serve as a valuable reference for fellow scholars embarking on future research. It is imperative to emphasize that model performance is intricately linked to variables such as tree species and planting density. Consequently, Figure 6 serves as an initial visual comparative reference, with further detailed deliberation remaining essential.

#### 4.5. Potential of Large Scale AGB Modeling

The current carbon sequestration potential of forests and their resilience against geographical and climatic changes are often assessed on the basis of AGB levels, which reflect forest productivity and resilience. In order to explore larger-scale forest AGB estimation, additional factors need to be further explored. Future work should further explore tree species diversity and heterogeneous forests for AGB prediction based on current experiments. In addition, taking all factors into consideration, the type of dataset closely determines the final performance of AGB estimation. Model precision is not the only factor in data source selection; the update frequency, range cover, and cost of the data are important as well [33]. Satellite imagery has a fixed revisit cycle, while low- and medium-resolution satellite imagery such as Landsat series images can be obtained for free.



The large-scale collection of ALS depends on the planning of data updated by the local forestry or surveying and mapping department. Although in this paper the model built using satellite imagery alone was less accurate compared to the best model built using ALS data, satellite imagery is widely available and inexpensive, and may even be available free of charge. With technological innovations, the cost of airborne point cloud data acquisition has recently been decreasing. ALS data production for ALTM 1210 acquisition was about USD 1100 per km<sup>2</sup> in 2004 [34] and about USD 500 per km<sup>2</sup> in 2010 [35]. Nevertheless, ALS data costs remain high compared to satellite imagery. Regarding the selection of image data sources for large-scale biomass modeling, further considerations can be made based on the results in Table 6. Our study is the first to use a combination of medium- and high-resolution satellite imagery and ALS data to investigate the cost and performance of multi-source data for AGB modeling. This study demonstrates the potential of using low-cost satellite data to provide a rough estimation of AGB nationally, and as such can guide future forest management decisions.

## 5. Conclusions

In this study, medium-resolution and high-resolution satellite imagery were combined with ALS data to estimate model performance and evaluate costs. First, TLS and ULS data were used to acquire individual tree parameters and sample AGB plots, replacing the traditional ground-measured data. The combination of ALS point clouds and satellite images was then used to obtain structure and spectral information of the forested areas. Satellite imagery, ALS, and fused features were extracted and filtered for AGB modeling using four regression methods. The most important part of this paper is the exploration of the performance of the different data sources, especially the performance of AGB estimation combined with medium-to-high resolution imagery and ALS data. The ALS model performed the best, followed by the GF2 model, while the LS8 model exhibited poorer performance. Among the different methods, the GF2-ALS model developed using SRM performed the best, with an  $R^2$  of 0.82. The cost needs to be considered during data selection as well; thus, remote sensing data with different costs were used to explore the potential of AGB estimation, including free and low-cost satellite data as well as more expensive airborne point cloud data. Using imagery alone as the data source, the GF2 model provided an increase in terms of  $R^2$  of 0.1 per square kilometer compared to the LS8 model, at an additional cost of 1 USD. When using fused data, the GF2-ALS data model  $R^2$  increased by 0.2, while the cost increased by USD 351 per square kilometer. Combining imagery with high spatial resolution and ALS data significantly improved the performance of the AGB model. Overall, it is necessary to comprehensively consider both cost and model performance for large-scale estimation of AGB.

**Author Contributions:** Methodology; writing—original draft preparation, W.F. Formal analysis; writing—review and editing, J.T. Writing—review and editing, T.K. Funding acquisition, B.Y. Visualization, F.L. Project administration, Z.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was jointly supported by the National Natural Science Foundation of China (No. 42130105) and the China Scholarship Council.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We wish to thank the anonymous reviewers and editors for their detailed comments and suggestions.

**Conflicts of Interest:** The authors declare no conflicts of interest.

Appendix A

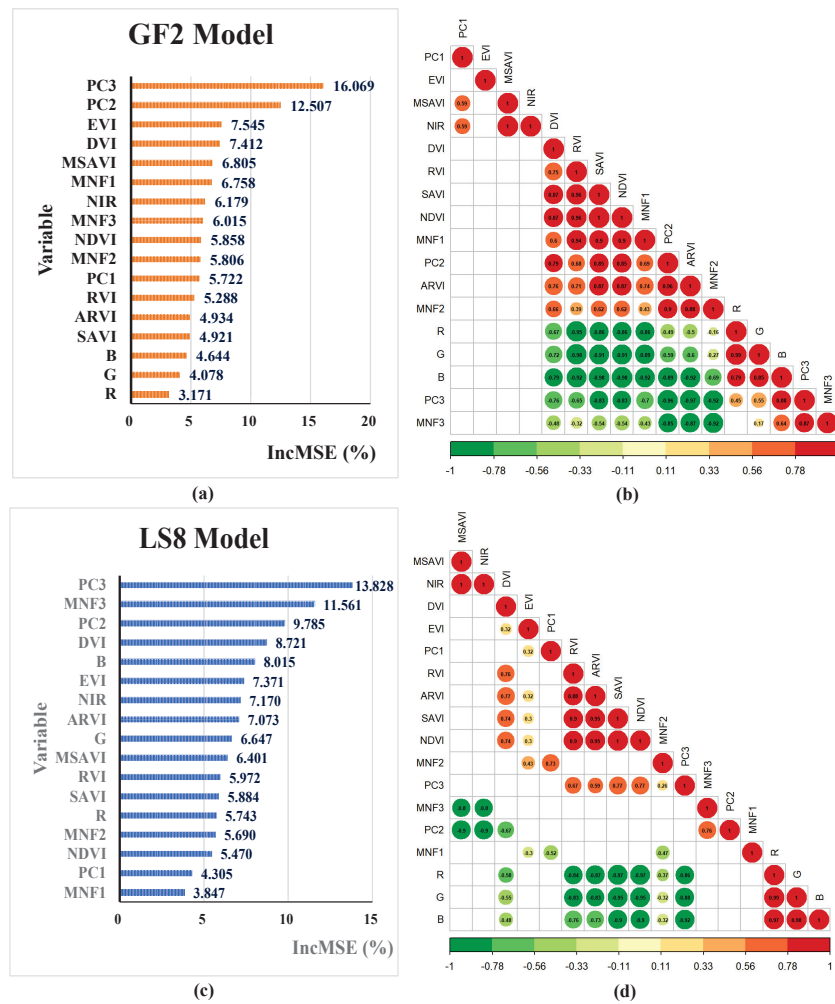


Figure A1. Satellite imagery variables selection. Importance ranking of the variables in (a) the GF2 model and (c) the LS8 model. Correlation analysis for (b) the GF2 model and (d) the LS8 model. Dark green and red indicate highly negative and positive correlations, respectively.

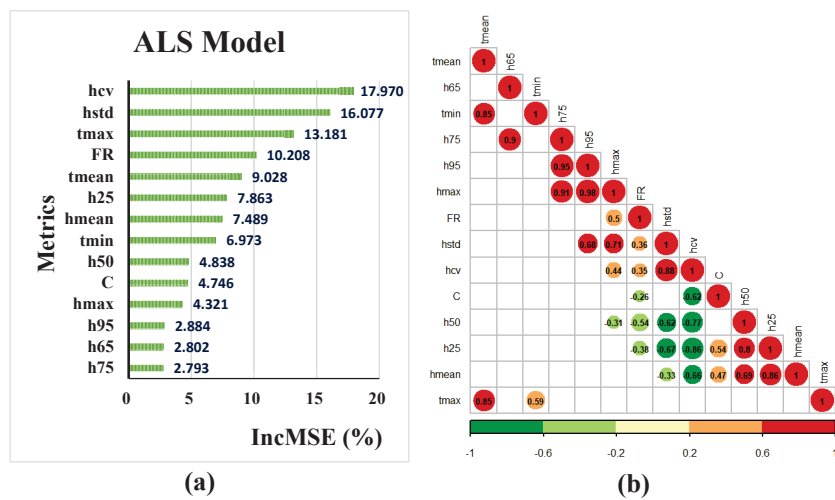
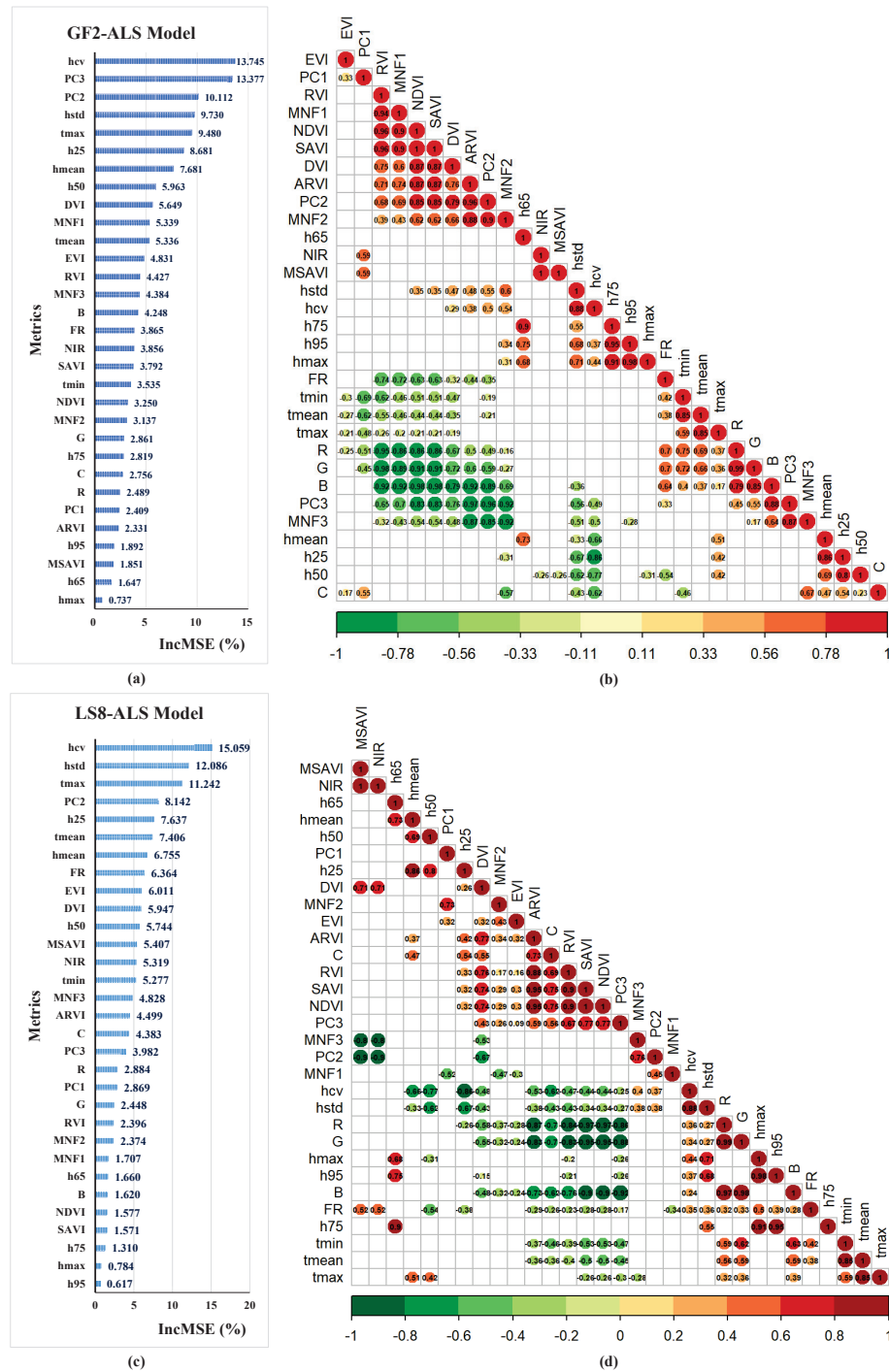


Figure A2. ALS variables selection. (a) Importance ranking of the variables in the ALS model. (b) Correlation analysis for the ALS model. Dark green and red indicate highly negative and positive correlations, respectively.



**Figure A3.** Fused data variables selection. (a) Importance ranking of variables in (a) the GF2-ALS model and (c) the LS8-ALS model. Correlation analysis for (b) the GF2-ALS model and (d) the LS8-ALS model. Dark green and red indicate highly negative positive correlations, respectively.

**References**

- Zhang, C.; Song, T.; Shi, R.; Hou, Z.; Wu, N.; Zhang, H.; Zhuo, W. Estimating the Forest Carbon Storage of Chongming Eco-Island, China, Using Multisource Remotely Sensed Data. *Remote Sens.* **2023**, *15*, 1575. [CrossRef]
- Zhu, J.; Hu, H.; Tao, S.; Chi, X.; Li, P.; Jiang, L.; Ji, C.; Zhu, J.; Tang, Z.; Pan, Y.; et al. Carbon stocks and changes of dead organic matter in China’s forests. *Nat. Commun.* **2017**, *8*, 151. [CrossRef]
- Lu, J.; Wang, H.; Qin, S.; Cao, L.; Pu, R.; Li, G.; Sun, J. Estimation of aboveground biomass of Robinia pseudoacacia forest in the Yellow River Delta based on UAV and Backpack LiDAR point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *86*, 102014. [CrossRef]

4. de Almeida, C.T.; Galvao, L.S.; Ometto, J.P.H.B.; Jacon, A.D.; de Souza Pereira, F.R.; Sato, L.Y.; Lopes, A.P.; de Alencastro Graça, P.M.L.; de Jesus Silva, C.V.; Ferreira-Ferreira, J.; et al. Combining LiDAR and hyperspectral data for aboveground biomass modeling in the Brazilian Amazon using different regression algorithms. *Remote Sens. Environ.* **2019**, *232*, 111323. [[CrossRef](#)]
5. Zhang, W.; Zhao, L.; Li, Y.; Shi, J.; Yan, M.; Ji, Y. Forest Above-Ground Biomass Inversion Using Optical and SAR Images Based on a Multi-Step Feature Optimized Inversion Model. *Remote Sens.* **2022**, *14*, 1608. [[CrossRef](#)]
6. Zhang, L.; Shao, Z.; Liu, J.; Cheng, Q. Deep learning based retrieval of forest aboveground biomass from combined LiDAR and landsat 8 data. *Remote Sens.* **2019**, *11*, 1459. [[CrossRef](#)]
7. Sankey, T.; Donager, J.; McVay, J.; Sankey, J.B. UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA. *Remote Sens. Environ. Interdiscip. J.* **2017**, *195*, 30–43. [[CrossRef](#)]
8. Ramachandran, N.; Saatchi, S.; Tebaldini, S.; d’Alessandro, M.M.; Dikshit, O. Mapping tropical forest aboveground biomass using airborne SAR tomography. *Sci. Rep.* **2023**, *13*, 6233. [[CrossRef](#)]
9. Zhang, Y.; Shao, Z. Assessing of urban vegetation biomass in combination with LiDAR and high-resolution remote sensing images. *Int. J. Remote Sens.* **2021**, *42*, 964–985. [[CrossRef](#)]
10. Cao, L.; Pan, J.; Li, R.; Li, J.; Li, Z. Integrating airborne LiDAR and optical data to estimate forest aboveground biomass in arid and semi-arid regions of China. *Remote Sens.* **2018**, *10*, 532. [[CrossRef](#)]
11. Momo Takoudjou, S.; Ploton, P.; Sonké, B.; Hackenberg, J.; Griffon, S.; De Coligny, F.; Kamdem, N.G.; Libalah, M.; Mofack, G.I.; Le Moguédec, G.; et al. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. *Methods Ecol. Evol.* **2018**, *9*, 905–916. [[CrossRef](#)]
12. Zhou, G.; Yin, G.; Tang, X. *Carbon Stocks in China’s Forest Ecosystems: A Biomass Equation*; Science Press: Beijing, China, 2018.
13. Dai, W.; Kan, H.; Tan, R.; Yang, B.; Guan, Q.; Zhu, N.; Xiao, W.; Dong, Z. Multisource forest point cloud registration with semantic-guided keypoints and robust RANSAC mechanisms. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *115*, 103105. [[CrossRef](#)]
14. Yang, B.; Dai, W.; Dong, Z.; Liu, Y. Automatic forest mapping at individual tree levels from terrestrial laser scanning point clouds with a hierarchical minimum cut method. *Remote Sens.* **2016**, *8*, 372. [[CrossRef](#)]
15. Dai, W.; Yang, B.; Liang, X.; Dong, Z.; Huang, R.; Wang, Y.; Li, W. Automated fusion of forest airborne and terrestrial point clouds through canopy density analysis. *ISPRS J. Photogramm. Remote Sens.* **2019**, *156*, 94–107. [[CrossRef](#)]
16. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
17. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
18. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with Erts. *NASA Spec. Publ.* **1974**, *351*, 309.
19. Birth, G.S.; McVey, G.R. Measuring the color of growing turf with a reflectance spectrophotometer 1. *Agron. J.* **1968**, *60*, 640–643. [[CrossRef](#)]
20. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
21. Huete, A.R.; A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
22. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [[CrossRef](#)]
23. Kaufman, Y.J.; Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 261–270. [[CrossRef](#)]
24. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-Validation. In *Encyclopedia of Database Systems*; Liu, L., Özsu, M.T., Eds.; Springer: Boston, MA, USA, 2009; pp. 532–538. [[CrossRef](#)]
25. Ioki, K.; Tsuyuki, S.; Hirata, Y.; Phua, M.H.; Wong, W.V.C.; Ling, Z.Y.; Saito, H.; Takao, G. Estimating above-ground biomass of tropical rainforest of different degradation levels in Northern Borneo using airborne LiDAR. *For. Ecol. Manag.* **2014**, *328*, 335–341. [[CrossRef](#)]
26. Han, H.; Wan, R.; Li, B. Estimating forest aboveground biomass using Gaofen-1 images, Sentinel-1 images, and machine learning algorithms: A case study of the Dabie Mountain Region, China. *Remote Sens.* **2021**, *14*, 176. [[CrossRef](#)]
27. Rana, P.; Gautam, B.; Tokola, T. Optimizing the number of training areas for modeling above-ground biomass with ALS and multispectral remote sensing in subtropical Nepal. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 52–62. [[CrossRef](#)]
28. Apollo Mapping. Apollo Mapping: The Image Hunters. 2022. Available online: <http://www.apollomapping.com/> (accessed on 10 August 2022).
29. Yang, Q.; Su, Y.; Hu, T.; Jin, S.; Liu, X.; Niu, C.; Liu, Z.; Kelly, M.; Wei, J.; Guo, Q. Allometry-based estimation of forest aboveground biomass combining LiDAR canopy height attributes and optical spectral indexes. *For. Ecosyst.* **2022**, *9*, 100059. [[CrossRef](#)]
30. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Del Frate, F.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58. [[CrossRef](#)]
31. Wang, C.; Nie, S.; Xi, X.; Luo, S.; Sun, X. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* **2016**, *9*, 11. [[CrossRef](#)]

32. Ghosh, S.M.; Behera, M.D. Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data-The superiority of deep learning over a semi-empirical model. *Comput. Geosci.* **2021**, *150*, 104737. [[CrossRef](#)]
33. Price, B.; Waser, L.T.; Wang, Z.; Marty, M.; Ginzler, C.; Zellweger, F. Predicting biomass dynamics at the national extent from digital aerial photogrammetry. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *90*, 102116. [[CrossRef](#)]
34. Eid, T.; Gobakken, T.; Næsset, E. Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scand. J. For. Res.* **2004**, *19*, 512–523. [[CrossRef](#)]
35. Montagni, A.; Corona, P.; Dalponte, M.; Gianelle, D.; Chirici, G.; Olsson, H. Airborne laser scanning of forest resources: An overview of research in Italy as a commentary case study. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 288–300. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.